

REMARKS

The above amendments to the above-captioned application along with the following remarks are being submitted as a response to the Official Action dated June 5, 2006. In view of the above amendments and the following remarks, the Examiner is respectfully requested to give due reconsideration to this application, to indicate the allowability of the claims, and to pass this case to issue.

Status of the Claims

Claims 1, 3-4 and 6-8 are under consideration in this application. Claims 2 and 5 are being cancelled without prejudice or disclaimer. Claims 1, 3-4 and 6 are being amended, as set forth in the above marked-up presentation of the claim amendments, in order to more particularly define and distinctly claim applicants' invention.

The claims are being amended to correct formal errors and/or to better recite or describe the features of the present invention as claimed. All the amendments to the claims are supported by the specification. Applicants hereby submit that no new matter is being introduced into the application through the submission of this response.

Formality Rejection

The Examiner rejected claims 1-8 under 35 U.S.C. §112, first paragraph, as failing to comply with the written description requirement. In particular, the Examiner has contended that the claims contain new matter since the specification describes a database which contains a nucleic acid sequence, but not an amino acid sequence.

Applicants respectfully contend that the disclosure uses a nucleic acid sequence AGCTAGCTAGCTAGCTAGCT (Fig. 2) only as an example for the base or amino acid sequence 201 (p. 4, lines 24-25; "*A user enters a base or amino acid sequence, such as a base sequence AGCT, for example*" p. 5, lines 23-24)". The base or amino acid sequence 201 can be any amino acid sequence including building units other than A, G, C, T as known to one skilled in the art and evidenced by the publications concurrently filed via the IDS.

In addition, by reading the recitation together with "of a gene or protein," the first database 105 which contains a base sequence or an amino acid sequence "of a gene or protein" actually contain only any nucleic acid sequences including building units of A, G, C, T, since they are "of a gene or protein."

Accordingly, the withdrawal of the outstanding informality rejection is in order, and is therefore respectfully solicited.

Prior Art Rejections

Claims 1, 3, 4 and 6 were rejected under 35 U.S.C. § 102(e) as being anticipated by US Pat. App. Pub. No. 2002/0168664 issued into US. Pat. No. 6,876,930 to Murray et al. (hereinafter “Murray”), and claims 2 and 5 were rejected under 35 U.S.C. § 103 as being unpatentable over Murray in view of US Patent No. 6,519,592 to Getchius (hereinafter “Getchius”). These rejections have been carefully considered, but are most respectfully traversed in view of the newly submitted claims, as more fully discussed below.

The method of calculating the frequency of appearance of a keyword of the invention (for example, the embodiment depicted in Fig. 8), as now recited in claim 1, uses a first database in which information about a base sequence or an amino acid sequence is stored and a second database in which document data is stored (for example, per record or per document; Figs. 2-3 show the structure of first and second documents, and “*each file of text data 106 contains identifier 202 for identifying document data*” p. 5, last paragraph). The method comprises: a first document extraction step (Steps 803+) for extracting a first document 106 (for example, “*a thesis describing the result of research into a particular base sequence*” p. 4, last paragraph) from said first database 105 (Fig. 2) which contains a base sequence or an amino acid sequence of a gene or protein of interest (e.g., “AGCT” p. 5, 2nd to last paragraph, 201 in Fig. 2, or 401 in Fig. 4) inputted by a user; an identifier extraction step (Steps 806+) for extracting an identifier 202 (e.g., P00005) identifying document data (e.g., Fig. 2: “As seen in the example”) in said first document 106 from said extracted first document 106 which contains the base sequence or the amino acid sequence 201; a second document extraction step (Step 807) for extracting a second document 108 (for example, “*the data describes the result of molecular-biological study into a gene or protein*” p. 5, 1st paragraph) from said second database 107 (Fig. 3) which contains said extracted identifier 301 (same as 202, e.g., P00005); an appearance frequency calculation step (Step 810) for sequentially reading keywords 503 (e.g., “axon midline choice point recognition” etc. listed in the last column of the table in Fig. 5) from a keyword table 110 (Fig. 5; “*The calculating unit 102 obtains keywords contained in the category table 110 in the third file system 109*” p. 6, lines 3-4) containing keywords of known functions or characteristics of genes or proteins from said first database, and for calculating a frequency of appearance of each of said keywords 501-503 by counting a number of second documents containing said keywords 503 in said extracted second documents 108 (“*Specifically, the number of files of extracted text data 108 in which each keyword appears or is used is calculated.*” p. 6, lines 5-6); and a

displaying step for displaying a frequency of appearance of each of said keywords 501-503 in a corresponding position in said keyword table 110 (e.g., Figs. 6-7). The keyword table 110 has a tree structure in which keywords 503 are stored such that they are classified according to categories 501, 502. The appearance frequency calculation step comprises a step for generating a frequency calculation result table (e.g., Fig. 7) of a tree structure, said table containing the frequency of appearance of a keyword 503 and the frequency of appearance of an upper-level category 501 to which the keyword belongs (“*In the illustrated example, the frequency of appearance of all of the keywords belonging to the category “cell recognition” is 196. This indicates that keywords belonging to the category “cell recognition” appear at least once in 196 files of the text data contained in the second file system 107.*” p. 6, last paragraph).

The invention recited in claim 4 is directed to a program embedded in a storage medium for causing a computer to carry out the keyword frequency calculation method of claim 1.

“*The user can thus learn the frequency of each keyword related to the sequence AGCT in the text data 108 in the second file system 107* (p. 6, 2nd paragraph),” so as to find out the functions or characteristics of the particular gene or protein (p. 1, lines 18-19; P. 3, 2nd paragraph).

In accordance with the invention, when a searcher wishes to know the functions or characteristics of a gene or protein with a particular sequence, the searcher can be provided with a list of keywords indicating the functions or characteristics of genes or proteins containing the particular sequence by entering the particular sequence information itself as a search key, the list showing the keywords in terms of the importance, or the frequency of appearance in document data (p. 3, 2nd paragraph). By viewing the keyword frequency table in Fig. 7, the research can know how much literature exists on the subject to know whether the subject has already been researched or studied enough so as to determine whether to select or change the research’s course (p. 3, lines 8-16).

Applicants respectfully contend that none of the cited references teaches or suggests “a second document extraction step for extracting a second document 108 from said second database 107 which contains the same identifier 202 (as contained in the said first document 106) of the base sequence or the amino acid sequence of a gene or protein of interest” or “an appearance frequency calculation step for sequentially reading keywords 501-503 from a keyword table 110 (Fig. 5) containing keywords of known functions or characteristics of genes or proteins from said first database 105 and for calculating a frequency of appearance

of each of said keywords 503 by counting a number of second documents containing said keywords in said extracted second document 108" according to the invention.

In contrast, Murray conducts research at the gene level ("*to provide further information about a particular DNA sequence to facilitate its identification as a candidate gene*" Abstract), without involving any "base sequence or an amino acid sequence" of a gene or protein of interest (which is shorter than a gene) as the invention. Murray also does not link genes/proteins of known functions/characteristics with the gene/protein of interest via the particular "base sequence or an amino acid sequence" contained in both the gene/protein of interest and the genes/proteins of known functions/characteristics.

In addition, Murray just either (1) uses "journal titles" as keywords to quantitatively score the times each of the journal titles was "*referenced in other journals*", or (2) adopts the quality scores listed in Table 1 of Murray (p. 10) corresponding to a plurality of journals ([0101]), in order to evaluate a given biological topic as a majority or a minority view model ([0099]). For example, an A-B model scored between 51-100 constitutes a majority view model, while an A-C-B model scored between 0-49 constitutes a minority view model.

In the (1) embodiment, Murray's "*journal titles*" for quantity counting are essentially different from the "*known functions or characteristics*" of genes or proteins, such as "axon midline choice point recognition" etc. as keywords for accounting according to the present invention. In addition, Murray's "*journal titles*" are used to count an article being referenced in other journals, i.e., a number counting *per journal*, rather than to count numbers of second documents including that keyword, i.e., a number-counting *per document*, according to the present invention.

In the (2) embodiment, Table 1 of Murray (p. 10) only shows quality scores corresponding to a plurality of journals, which are different from the keyword table 110 (Fig. 5) listing the quantity/number of documents containing the keywords of the invention. Murray's quality scores are determined by reputation. For example, "Science", "Nature", the "Journal of Biological Chemistry", or the "Journal of Clinical Investigations" subject to a recognized, rigorous peer-review process receive 100, 100, 92, 90 points respectively, whereas the "Journal of Antisense Research or Experimental Cell Research" only receives 10 points. Murray then uses the quality scores as weighting ([0102]) to adjust the quantity scores, i.e., a number of times being referenced in other journals in the (1) embodiment. Again, the adjusted quantity scores of Murray are essentially different from the documents counts of the present invention, since Murray counts *per journal*, rather than automatically

and mechanically counting *per document* (not involving human opinion), as does the present invention.

Getchius was relied upon by the Examiner (p. 6, 4th paragraph of the outstanding Office Action) to teach using a tree structure for the keyword table as recited in claims 2 and 5 of the invention which are now incorporated in to claims 1 and 4. However, Getchius is applied to a business listing (Abstract; e.g., *“restaurant” may be stored in a tree that includes the sub-category of “ethnic restaurant,” which may further include the sub-category “Greek restaurant.””* col. 33, lines 19-21), rather than any “known functions or characteristics of genes or proteins” as the keywords in the keyword table 110 of the invention.

To adequately traverse the Examiner’s assertions, Applicants respectfully contend that one skilled in the art would not be motivated to “display a frequency of appearance of each of said keywords in a corresponding position in said tree-structured keyword table 110” as claimed by the Applicants based on the Getchius’ tree-structured business list, except by using Applicants’ method as a blueprint to adopt “known functions or characteristics of genes or proteins from said first database 105” as keywords. Getchius’ tree-structured business list simply did not constitute common knowledge in the art of gene/protein research, especially regarding adopting “known functions or characteristics of genes or proteins from said first database 105” as keywords in a tree structure as shown in Fig. 5 of the invention.

Moreover, the proposed combination of Getchius’ tree-structured business list with Fig. 7 of Murray would totally destroy the intended purpose for Fig. 7: annotating functional information for genes per reference, i.e., *“a single published article* (p. 12, left col. line 7-8)”. Murray’s Fig. 7 is *“a table of the annotation summary (“AnnotationSummary” 160) including the sequence name (“SeqFile” 160-a), best hits (“BestHits” 160-b) which refers to the “DNAsequence” table 130 (“BestBlastnGID” 130-f), a link to the “Function” table 170 (“Function” 160-c), a link to the “Role” table 180 (“Role” 160-d), a link to the “Evidence” table 190 (“Evidence” 160-e). The Function 170, Role 180 and Evidence 190 tables contain many attributes which all refer to individual References (“Reference” table 200)([0106])”*. Murray’s functional information table Fig. 7 simply does not have any quantitative relationships among the root, the trunk, and the branches as Getchius’ tree-structured business list. One skilled in the art would not be motivated to combine the teachings in Getchius and Murray in the manner suggested by the Examiner since the resulting change in the organization of Fig. 7 in Murray will contradict its intended purpose. It is well established that a rejection based on a principle that contracts the teachings of the cited references is also improper.

Applicants will further contend that the combination of references used by the

Examiner merely consists of selecting bits and pieces from each reference, and then combining those bits and pieces using knowledge or hindsight gleaned from the disclosure of the present invention as a guide to support the combination. The well established rule of law is that each prior art reference must be evaluated as an entirety, and that the entire prior art must be considered as a whole,” *Panduit Corp. v. Dennison Mfg. Co.*, 227 USPQ 337, 344 (Fed. Cir. 1985). See *Para-Ordinance Mfg. Inc. v. SGS Importers Intl.*, Inc., 73 F.3d 1085, 37 USPQ2d 1237 (Fed. Cir. 1995) (“Obviousness may not be established using hindsight or in view of the teachings or suggestions of the inventor.”).

Even if, arguendo, one skilled in the art were motivated to combine the teachings in Murray and Getchius as suggested by the Examiner, such combined teachings would still fall short in fully meeting the Applicants' claimed invention as set forth in claims 1 and 4, since neither reference teaches adopting “a second document extraction step for extracting second document 108 from said second database 107 which contains the same identifier 202 (as contained in the said first document 106) of the base sequence or the amino acid sequence of a gene or protein of interest” or “an appearance frequency calculation step for sequentially reading keywords 501-503 from a keyword table 110 (Fig. 5) containing keywords of known functions or characteristics of genes or proteins from said first database 105 and for calculating a frequency of appearance of each of said keywords 503 by counting a number of second documents containing said keywords in said extracted second document 108”, as in the invention.

Applicants contend that neither Murray, Getchius, nor their combinations teaches or discloses each and every feature of the present invention as recited in independent claims 1 and 4. As such, the present invention as now claimed is distinguishable and thereby allowable over the rejections raised in the Office Action. The withdrawal of the outstanding prior art rejections is in order, and is respectfully solicited.

Conclusion

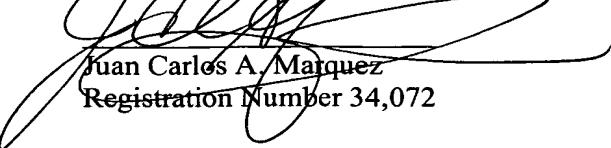
In view of all the above, clear and distinct differences as discussed exist between the present invention as now claimed and the prior art reference upon which the rejections in the Office Action rely, Applicants respectfully contend that the prior art references cannot anticipate the present invention or render the present invention obvious. Rather, the present invention as a whole is distinguishable, and thereby allowable over the prior art.

Favorable reconsideration of this application is respectfully solicited. Should there be any outstanding issues requiring discussion that would further the prosecution and allowance

of the above-captioned application, the Examiner is invited to contact the Applicants' undersigned representative at the address and telephone number indicated below.

Respectfully submitted,

Stanley P. Fisher
Registration Number 24,344


Juan Carlos A. Marquez
Registration Number 34,072

REED SMITH LLP
3110 Fairview Park Drive, Suite 1400
Falls Church, Virginia 22042
(703) 641-4200

September 5, 2006

SPF/JCM/JT